On Optimal Solutions to Two-Block H^∞ Problems *

Babak Hassibi and Thomas Kailath

Information Systems Laboratory

Stanford University

September 16, 1997

Abstract

In this paper we obtain a new formula for the minimum achievable disturbance attenuation in two-block H^∞ problems. This new formula has the same structure as the optimal H^∞ norm formula for noncausal problems, except that doubly-infinite (so-called Laurent) operators must be replaced by semi-infinite (so-called Toeplitz) operators. The benefit of the new formula is that it allows us to find explicit expressions for the optimal H^∞ norm in several important cases: the equalization problem (or its dual, the tracking problem), and the problem of filtering signals in additive noise. Furthermore, it leads us to the concepts of "worst-case non-estimability", corresponding to when causal filters cannot reduce the H^∞ norms from their a priori values, and "worst-case complete estimability", corresponding to when causal filters offer the same H^∞ performance as noncausal ones. We also obtain an explicit characterization of worst-case non-estimability and study the consequences to the problem of equalization with finite delay.

1 Introduction

The formula for the minimum achievable disturbance attenuation in two-block H^∞ problems (denoted hereafter by γ_c) was obtained by Verma and Jonckheere [1] and Feintuch and Francis

*This work was supported in part by DARPA through the Department of Air Force under contract F49620-95-1-0525-P00001 and by the Joint Service Electronics Program at Stanford under contract DAAH04-94-G-0058-P00003.

[2] in the mid 1980's. In these works γ_c was described as the spectral radius of a mixed Toeplitz-plus-Hankel operator. Despite the elegance of this result, little physical insight into the properties of the two-block problem have been obtained in this framework. Moreover, the explicit computation of the spectral radius of the mixed Toeplitz-plus-Hankel operator has been superseded by state-space Riccati-based approaches that compute this quantity only implicitly (see *e.g.*, [4]).

In this paper we attempt to show that much insight into the two-block H^∞ problem can be obtained by studying the minimum achievable disturbance attenuation. The main result in this attempt is a new formula for γ_c which is, in our view, simpler than the mixed Toeplitz-plus-Hankel spectral radius formula, and which retains the same structure as the formula for the minimum achievable disturbance attenuation in noncausal two-block H^∞ problems. To demonstrate this fact, we give a complete analysis of γ_c for the important problems of equalization and of filtering signals in additive noise. This study reveals that for minimum phase systems causal equalizers have the same H^∞ performance as non-causal ones, whereas for non-minimum phase systems causal equalizers cannot reduce further attenuate the disturbances from their a priori values. For the problem of filtering signals in additive noise, the study reveals that causal filters have the same H^∞ performance as non-causal ones.

While such a complete analysis for general two block problems is not currently available (primarily due to the fact that simple frequency-domain characterizations of γ_c currently do not exist), we do introduce two concepts that we believe are of considerable importance. The first is "worst-case complete estimability" which corresponds to when causal estimators have the same H^∞ performance as non-causal ones, and roughly speaking represents an *easy* estimation problem. The second is "worst-case non-estimability" which corresponds to when causal estimators cannot reduce γ_c from their a priori bounds, and roughly speaking represents a *difficult* estimation problem. These concepts are important since in estimation (and control) one would like to set up problems that are close to worst-case complete estimable and that avoid worst-case non-estimability. Although we have not been able to give a simple characterization of worst-case complete estimability, we do give one for worst-case non-estimability and demonstrate its merits by studying the problem of equalization with finite delay.

The remainder of the paper is organized as follows. We introduce our notation in Sec. 1.1 and state the two-block H^∞ problem in Sec. 2. The non-causal solution to this problem is given in Sec. 2.1. The main result of the paper is in Sec. 3 which gives a new formula for the minimum achievable disturbance attenuation, γ_c. The special cases of equalization and filtering signals in additive noise are studied in Secs. 4.1 and 4.2, respectively. The concepts of worst-case complete estimability and worst-case non-estimability are introduced in Sec. 5, the characterization of worst-case non-estimability in Sec. 5.1, and the application to equalization with finite delay in Sec. 5.2. The conclusion is given in Sec. 6.

1.1 Notation

In this paper we shall often deal with time-invariant operators that map $l^{2,m}$ to $l^{2,p}$, *i.e.*, the space of square-summable sequences of m-vectors to the space of square-summable sequences of p-vectors, according to the rule

$$y_i = \sum_{j=-\infty}^{\infty} T_{i-j} u_j, \tag{1}$$

where $u = \{u_i\} \in l^{2,m}$ and $y = \{y_i\} \in l^{2,p}$. Note that the above equation can also be written as

$$\begin{bmatrix} \vdots \\ y_{-1} \\ y_0 \\ y_1 \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \ddots & \ddots & \vdots & \vdots & \\ \ddots & T_0 & T_{-1} & T_{-2} & \dots \\ \dots & T_1 & T_0 & T_{-1} & \dots \\ \dots & T_2 & T_1 & T_0 & \vdots \\ & \vdots & \vdots & \ddots & \ddots \end{bmatrix}}_{\triangleq \mathcal{T}} \begin{bmatrix} \vdots \\ u_{-1} \\ u_0 \\ u_1 \\ \vdots \end{bmatrix}. \tag{2}$$

Another characterization of the operator $\mathcal{T}$ is through its z-transform,

$$T(z) = \sum_{i=-\infty}^{\infty} T_i z^{-i}, \tag{3}$$

which, since $\mathcal{T}$ maps $l^{2,m}$ to $l^{2,p}$, converges absolutely on an annulus containing the unit circle, $|z| = 1$. This implies that the Fourier transform $T(e^{j\omega})$ is well-defined and bounded for all $\omega \in [0, 2\pi]$.

We shall also often find it useful to partition the input and output sequences u and y into the their past, $u_- \triangleq \{u_i, i < 0\}$ and $y_- \triangleq \{y_i, i < 0\}$, and present and future, $u_+ \triangleq \{u_i, i \geq 0\}$ and $y_+ \triangleq \{y_i, i \geq 0\}$, components. We shall also denote the corresponding spaces of semi-infinite sequences by $l_-^{2,m}$, $l_-^{2,p}$ and $l_+^{2,m}$, $l_+^{2,p}$, respectively. With this partitioning of the input and output spaces, the operator $\mathcal{T}$ can be partitioned as follows:

$$\mathcal{T} = \left[\begin{array}{c|c} \mathcal{T}_- & \mathcal{T}_A \\ \hline \mathcal{T}_H & \mathcal{T}_+ \end{array} \right], \tag{4}$$

where

$$\left\{ \begin{array}{cc} \mathcal{T}_- \triangleq \begin{bmatrix} \ddots & \ddots & \vdots & \vdots \\ \ddots & T_0 & T_{-1} & T_{-2} \\ \cdots & T_1 & T_0 & T_{-1} \\ \cdots & T_2 & T_1 & T_0 \end{bmatrix} & \mathcal{T}_A \triangleq \begin{bmatrix} \vdots & \vdots & \ddots \\ T_{-2} & T_{-3} & \cdots \\ T_{-1} & T_{-2} & \cdots \end{bmatrix} \\ \mathcal{T}_H \triangleq \begin{bmatrix} \cdots & T_2 & T_1 \\ \cdots & T_3 & T_2 \\ \ddots & \vdots & \vdots \end{bmatrix} & \mathcal{T}_+ \triangleq \begin{bmatrix} T_0 & T_{-1} & T_{-2} & \cdots \\ T_1 & T_0 & T_{-1} & \cdots \\ T_2 & T_1 & T_0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix} \end{array} \right. . \tag{5}$$

The operator $\mathcal{T}_-$ maps $l_-^{2,m}$ to $l_-^{2,p}$, *i.e.*, past inputs to past outputs, and is called a *Toeplitz* operator, whereas the operator $\mathcal{T}_H$ maps $l_-^{2,m}$ to $l_+^{2,p}$, *i.e.*, past inputs to present and future outputs, and is called a *Hankel* operator. Similar remarks apply to $\mathcal{T}_+$ and $\mathcal{T}_A$, though we shall not be concerned with such operators in this paper. The doubly-infinite operator $\mathcal{T}$, however, is referred to as a *Laurent* operator.

We will also be interested in the induced 2-norm, or so-called H^∞ norm, of such operators which is defined as

$$\|\mathcal{T}\|_\infty \triangleq \sup_{u \neq 0 \in l^{2,m}} \frac{\|\mathcal{T}u\|_2}{\|u\|_2} \quad \text{and} \quad \|\mathcal{T}_-\|_\infty \triangleq \sup_{u_- \neq 0 \in l_-^{2,m}} \frac{\|\mathcal{T}_- u_-\|_2}{\|u_-\|_2} \tag{6}$$

where we have used $\|a\|^2 = \sum_i a_i^* a_i$. Moreover, it turns out that, since Laurent and Toeplitz operators are time-invariant, there is a very simple frequency-domain characterization of their

H^∞ norms:

$$\|\mathcal{T}\|_\infty = \sup_{\omega\in[0,2\pi]} \bar{\sigma}\left[T(e^{j\omega})\right] \quad \text{and} \quad \|\mathcal{T}_-\|_\infty = \sup_{\omega\in[0,2\pi]} \bar{\sigma}\left[T(e^{j\omega})\right], \tag{7}$$

where $\bar{\sigma}(\cdot)$ denotes the maximum singular value of its argument. [Note that $\|\mathcal{T}\|_\infty = \|\mathcal{T}_-\|_\infty$.]

Finally, we should mention that the operators $\mathcal{T}$ and $\mathcal{T}_-$ will be called causal if $T_i = 0$, for all $i < 0$, or, in other words, if $\mathcal{T}_A = 0$.

2 The Two-Block Problem

Consider the following "two-block" operator,

$$\mathcal{T}_\mathcal{K} = \left[\begin{array}{cc} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{array}\right], \tag{8}$$

where $\mathcal{L}$ and $\mathcal{H}$ are causal Laurent operators. The operator $\mathcal{K}$ is also Laurent, though not necessarily causal. The two block H^∞ problem can thus be formulated as follows.

Problem 1 (Two-Block H^∞ Problem) *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$.*

(a) Find γ_s, where

$$\gamma_s \triangleq \inf_{\mathcal{K}} \left\| \left[\begin{array}{cc} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{array}\right] \right\|_\infty . \tag{9}$$

(b) Find γ_c, where

$$\gamma_c \triangleq \inf_{\text{causal } \mathcal{K}} \left\| \left[\begin{array}{cc} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{array}\right] \right\|_\infty . \tag{10}$$

Note that in Problem 1-(a) there is no restriction on the Laurent operator $\mathcal{K}$, whereas in Problem 1-(b) the Laurent operator is restricted to being causal. It is thus clear that

$$\gamma_s \leq \gamma_c. \tag{11}$$

Many estimation problems lead to Problem 1. To this end, consider Fig. 1 where $\mathcal{H}$ and $\mathcal{L}$ are known causal linear time-invariant operators (or simply, causal LTI systems), the sequences $\{u_i\}$ and $\{v_i\}$ are unknown sequences, $\{y_i\}$ is the known observations sequence, and $\{s_i\}$ is the unobservable desrired sequence we wish to estimate. The goal in estimation is to appropriatley design the (so-called) estimator $\mathcal{K}$ that provides the estimates $\{\hat{s}_i\}$ based on the observations $\{y_i\}$. When $\mathcal{K}$ is a non-causal operator the estimation problem is referred to as a *smoothing* problem since the estimator has access to future observations, whereas when $\mathcal{K}$ is a causal operator the estimation problem is referred to as a *filtering* problem since the estimator does not have access to future observations.

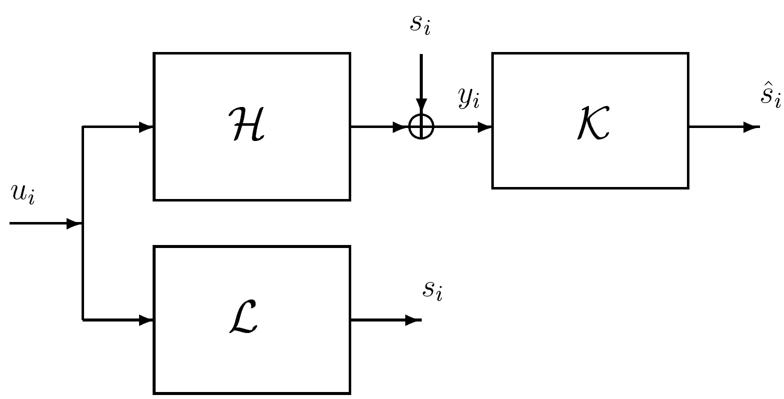

Figure 1: *A general estimation problem.*

The behavior of any estimator $\mathcal{K}$ may be captured by the induced transfer operator that maps the unknown disturbances $\{u_j\}$ and $\{v_j\}$ to the estimation errors $\tilde{s} \triangleq \{\tilde{s}\}$. However, from Fig. 1 it straightforward to see that this transfer operator is given by $\mathcal{T}_{\mathcal{K}}$, *i.e.*,

$$\mathcal{T}_{\mathcal{K}} : \begin{bmatrix} u \\ v \end{bmatrix} \rightarrow \tilde{s}. \tag{12}$$

In estimation the goal is to make the transfer operator $\mathcal{T}_{\mathcal{K}}$ small in some sense. In Problem 1 we have proposed to make $\mathcal{T}_{\mathcal{K}}$ small in the sense of its H^∞ norm. This will have the effect of minimizing the maximum energy gain from the disturbances to the estimation errors, *i.e.*,

$$\inf_{\mathcal{K}} \sup_{u,v \neq \in l^2} \frac{\|\tilde{s}\|_2^2}{\|u\|_2^2 + \|v\|_2^2} \quad \text{and} \quad \inf_{\text{causal } \mathcal{K}} \sup_{u,v \neq \in l^2} \frac{\|\tilde{s}\|_2^2}{\|u\|_2^2 + \|v\|_2^2}, \tag{13}$$

which is also the reason why γ_s and γ_c are referred to as the minimum disturbance attenuations.

Many full information control problems also lead to the two-block problem of Problem 1. However, since they are essentially the dual of the estimation problem just mentioned, we shall not further consider them here.

Finally, we should mention that there is no loss of generality in assuming that $\mathcal{H}$ and $\mathcal{L}$ are causal. When $\mathcal{H}$ and $\mathcal{L}$ are noncausal, the two-block problem can be readily replaced by an equivalent two-block problem where the $\mathcal{H}$ and $\mathcal{L}$ are indeed causal.

2.1 The Non-Causal Solution

Finding an expression for the minimum disturbance attenuation when $\mathcal{K}$ is not restricted to be causal is quite straightforward and only requires a "completion of squares" argument.

Theorem 1 (Non-Causal Optimal H^∞ Norm) *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$ and suppose we would like to solve*

$$\gamma_s = \inf_{\mathcal{K}} \left\| \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty .$$

Then we have

$$\gamma_s^2 = \left\| \mathcal{L} \left(I + \mathcal{H}^*\mathcal{H} \right)^{-1} \mathcal{L}^* \right\|_\infty = \sup_{\omega \in [0,2\pi]} \bar{\sigma} \left[L(e^{j\omega}) \left(I + H^*(e^{j\omega}) H(e^{j\omega}) \right)^{-1} L^*(e^{j\omega}) \right]. \tag{14}$$

Proof: Note that we may write

$$\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^* = (\mathcal{L} - \mathcal{K}\mathcal{H})(\mathcal{L} - \mathcal{K}\mathcal{H})^* + \mathcal{K}\mathcal{K}^*,$$

so that after a completion of squares,

$$\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^* = \left(\mathcal{K} - \mathcal{L}\mathcal{H}^*(I + \mathcal{H}\mathcal{H}^*)^{-1} \right) (I + \mathcal{H}\mathcal{H}^*) \left(\mathcal{K} - \mathcal{L}\mathcal{H}^*(I + \mathcal{H}\mathcal{H}^*)^{-1} \right)^* + \mathcal{L} \left(I + \mathcal{H}^*\mathcal{H} \right)^{-1} \mathcal{L}^*. \tag{15}$$

Since the second term on the RHS is independent of $\mathcal{K}$, the transfer operator $\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*$ can, in fact, be minimized by setting the first term equal to zero (take $\mathcal{K} = \mathcal{L}\mathcal{H}^*(I + \mathcal{H}\mathcal{H}^*)^{-1}$). This leads to $\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^* = \mathcal{L} \left(I + \mathcal{H}^*\mathcal{H} \right)^{-1} \mathcal{L}^*$, and hence the desired result.

■

3 Optimal H^∞ Norm in the Causal Case

The main result of this paper has to do with a new expression for γ_c, which is given below.

Theorem 2 (Causal Optimal H^∞ Norm) *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$ and suppose we would like to solve*

$$\gamma_c = \inf_{\text{causal}\mathcal{K}} \left\| \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty .$$

Then we have

$$\gamma_c^2 = \left\| \mathcal{L}_- \left(I + \mathcal{H}_-^* \mathcal{H}_- \right)^{-1} \mathcal{L}_-^* \right\|_\infty . \tag{16}$$

Note that the expressions for the minimum disturbance attenuation in the non-causal and causal cases (of Theorems 1 and 2) have the exact same structure! The only difference between the two is that the doubly-infinite Laurent operators $\mathcal{L}$ and $\mathcal{H}$ must be replaced by the semi-infinite Toeplitz operators $\mathcal{L}_-$ and $\mathcal{H}_-$. This fact has the following interpretation: the minimum disturbance attenuation in a doubly-infinite causal estimation problem is the same as the the minimum disturbance attenuation in the semi-infinite *non-causal* problem.

The reason why we were able to give a simple frequency-domain expression for γ_s in Theorem 1 is that products and inverses of Laurent operators are themselves Laurent operators. Thus $\mathcal{L}\left(I + \mathcal{H}^*\mathcal{H}\right)^{-1} \mathcal{L}^*$ is a Laurent operator and Eq. (14) readily follows. However, products and inverses of Toeplitz operators are not necessarily Toeplitz (the reader may want to check, for example, that $\mathcal{H}_-^*\mathcal{H}_-$ is not Toeplitz). Thus $\mathcal{L}_- \left(I + \mathcal{H}_-^*\mathcal{H}_-\right)^{-1} \mathcal{L}_-^*$ is not necessarily a Toeplitz operator, and so simple frequency-domain formulas for γ_c in Eq. (16) cannot generally be given.

However, some further characterization of $\mathcal{L}_- \left(I + \mathcal{H}_-^*\mathcal{H}_-\right)^{-1} \mathcal{L}_-^*$ can indeed be given. To this end, let us define

$$\mathcal{E} \triangleq \mathcal{L}(I + \mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^* = \left[\begin{array}{c|c} \mathcal{E}_- & \mathcal{E}_H^* \\ \hline \mathcal{E}_H & \mathcal{E}_+ \end{array} \right], \tag{17}$$

and note that

$$\mathcal{E} = \mathcal{L}\mathcal{L}^* - \mathcal{L}\mathcal{H}^*(I + \mathcal{H}\mathcal{H}^*)^{-1}\mathcal{H}\mathcal{L}^*. \tag{18}$$

Now, if we further define the spectral factorization

$$\Delta\Delta^* = I + \mathcal{H}\mathcal{H}^*, \tag{19}$$

with Δ causal and causally invertible, and the operator

$$\mathcal{P} \triangleq \Delta^{-1}\mathcal{H}\mathcal{L}^* = \left[\begin{array}{c|c} \mathcal{P}_- & \mathcal{P}_A \\ \hline \mathcal{P}_H & \mathcal{P}_+ \end{array}\right], \tag{20}$$

then using (18) we can identify the Toeplitz part of $\mathcal{E}$ as

$$\mathcal{E}_- = (\mathcal{L}\mathcal{L}^*)_- - (\mathcal{P}^*\mathcal{P})_- = \mathcal{L}_-\mathcal{L}_-^* - \mathcal{P}_-^*\mathcal{P}_- - \mathcal{P}_H^*\mathcal{P}_H. \tag{21}$$

But since

$$\mathcal{P} = \left[\begin{array}{c|c} \Delta_-^{-1} & 0 \\ \hline -\Delta_+^{-1}\Delta_H\Delta_-^{-1} & \Delta_+^{-1} \end{array}\right] \left[\begin{array}{c|c} \mathcal{H}_- & 0 \\ \hline \mathcal{H}_H & \mathcal{H}_+ \end{array}\right] \left[\begin{array}{c|c} \mathcal{L}_-^* & \mathcal{L}_H^* \\ \hline 0 & \mathcal{L}_+^* \end{array}\right],$$

we have

$$\mathcal{P}_- = \Delta_-^{-1}\mathcal{H}_-\mathcal{L}_-^*. \tag{22}$$

This therefore implies that $\mathcal{E}_-$ can be written as

$$\mathcal{E}_- = \mathcal{L}_-\mathcal{L}_-^* - \mathcal{L}_-\mathcal{H}_-^*\Delta_-^{-*}\Delta_-^{-1}\mathcal{H}_-\mathcal{L}_-^* - \mathcal{P}_H^*\mathcal{P}_H. \tag{23}$$

But since

$$\left[\begin{array}{c|c} \Delta_- & 0 \\ \hline \Delta_H & \Delta_+ \end{array}\right] \left[\begin{array}{c|c} \Delta_-^* & \Delta_H^* \\ \hline 0 & \Delta_+^* \end{array}\right] = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}\right] + \left[\begin{array}{c|c} \mathcal{H}_- & 0 \\ \hline \mathcal{H}_H & \mathcal{H}_+ \end{array}\right] \left[\begin{array}{c|c} \mathcal{H}_-^* & \mathcal{H}_H^* \\ \hline 0 & \mathcal{H}_+^* \end{array}\right],$$

we have

$$\Delta_-\Delta_-^* = I + \mathcal{H}_-\mathcal{H}_-^*,$$

so that

$$\mathcal{E}_- = \mathcal{L}_-\mathcal{L}_-^* - \mathcal{L}_-\mathcal{H}_-^*(I + \mathcal{H}_-\mathcal{H}_-^*)^{-1}\mathcal{H}_-\mathcal{L}_-^* - \mathcal{P}_H^*\mathcal{P}_H = \mathcal{L}_-(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^* - \mathcal{P}_H^*\mathcal{P}_H. \tag{24}$$

We thus finally have

$$\mathcal{L}_-(I+\mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^* = \mathcal{E}_- + \mathcal{P}_H^*\mathcal{P}_H, \tag{25}$$

which recovers a wellknown result of Verma and Jonckheere [1] and Feintuch and Francis [2] that states that the minimum disturbance attenuation is given by the spectral radius of the so-called mixed Toeplitz-plus-Hankel operator $\mathcal{E}_- + \mathcal{P}_H^*\mathcal{P}_H$. We can summarize this result in the following theorem.

Theorem 3 (Mixed Toeplitz-Plus-Hankel Operator) *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$. Then we have*

$$\gamma_c^2 = \|\mathcal{E}_- + \mathcal{P}_H^*\mathcal{P}_H\|_\infty . \tag{26}$$

where we have defined

$$\mathcal{E} = \mathcal{L}(I+\mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*,$$

and

$$\mathcal{P} = \Delta^{-1}\mathcal{H}\mathcal{L}^* \quad , \quad \Delta\Delta^* = I + \mathcal{H}\mathcal{H}^*$$

with Δ causal and causally invertible.

Eq. (25) is significant since it shows that the operator $\mathcal{L}_-(I+\mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^*$ is generally not Toeplitz, and, in fact, that it differs from the Toeplitz operator $\mathcal{E}_-$ by the amount of $\mathcal{P}_H^*\mathcal{P}_H$. When the operators $\mathcal{L}$ and $\mathcal{H}$ arise from a finite-dimensional state-space model, *i.e.*, when the corresponding $L(z)$ and $H(z)$ are rational of McMillan degree n, say, then the Hankel operator $\mathcal{P}_H$ has finite rank n. [Thus $\mathcal{L}_-(I+\mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^*$ differs from a Toeplitz operator by a finite rank amount.]

Note that $\gamma_s^2 = \|\mathcal{E}_-\|_\infty$. Therefore (26) shows, as expected, that $\gamma_c \geq \gamma_s$, and that the increase in the minimum disturbance attenuation depends on the Hankel operator $\mathcal{P}_H$.

Although Eq. (26) is an intriguing result, it has not proven to be a very useful tool in analyzing the behaviour of γ_c with respect to $\mathcal{L}$ and $\mathcal{H}$. The main reason is that simple frequency domain formulas for the spectral radius of mixed Toeplitz-plus-Hankel operators are not generally available [3]. [Currently, for operators arising from finite-dimensional state-space models, γ_c is computed by iteratively solving certain algebraic Riccati equations [4].] Another reason is that the dependence of the Hankel operator $\mathcal{P}_H$ on the original $\mathcal{L}$ and $\mathcal{H}$ is quite complicated.

However, we shall in the remainder of this paper see that our new formula (16) allows us to draw various qualitative and quantitative conclusions about γ_c. The main reasons for this are that in the new formula the dependence of γ_c on $\mathcal{L}$ and $\mathcal{H}$ is much simpler, and that the formula retains the same structure as in the non-causal case of Eq. (14).

To end this section we shall give a proof of Theorem 2. Although it is possible to give a direct proof of this theorem, we shall instead prove Theorem 3. The proof of Theorem 2 then follows from the identity (25).

Proof of Theorems 2 and 3: As just stated, we need only prove Theorem 3. To this end, choose an arbitrary γ. Then using (15) we may write

$$\gamma^2 I - \mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K} = \gamma^2 I - \mathcal{L}(I+\mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^* + (\mathcal{K}\Delta - \mathcal{L}\mathcal{H}^*\Delta^{-*})(\mathcal{K}\Delta - \mathcal{L}\mathcal{H}^*\Delta^{-*})^*.$$

Now if γ is such that $\gamma > \gamma_s$, since $\gamma_s = \|\mathcal{L}(I+\mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*\|_\infty$, we can define the spectral factorization

$$\mathcal{S}\mathcal{S}^* = \gamma^2 I - \mathcal{L}(I+\mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^* = \gamma^2 I - \mathcal{E} > 0,$$

with $\mathcal{S}$ causal and causally invertible. We therefore have

$$\gamma^2 I - \mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K} = \mathcal{S}\left[I - (\mathcal{S}^{-1}\mathcal{K}\Delta - \mathcal{S}^{-1}\mathcal{L}\mathcal{H}^*\Delta^{-*})(\mathcal{S}^{-1}\mathcal{K}\Delta - \mathcal{S}^{-1}\mathcal{L}\mathcal{H}^*\Delta^{-*})^*\right].$$

The above equation implies that $\gamma^2 I - \mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K} > 0$, and $\gamma > \gamma_c$, if, and only if, a causal $\mathcal{K}$ can be chosen such that

$$(\mathcal{S}^{-1}\mathcal{K}\Delta - \mathcal{S}^{-1}\mathcal{L}\mathcal{H}^*\Delta^{-*})(\mathcal{S}^{-1}\mathcal{K}\Delta - \mathcal{S}^{-1}\mathcal{L}\mathcal{H}^*\Delta^{-*})^* < I.$$

But since $\mathcal{S}^{-1}\mathcal{K}\Delta$ is causal (and so are $\mathcal{S}$ and Δ^{-1}), due to Nehari's Theorem [5], this is possible if, and only if, the Hankel operator of $\mathcal{T} \triangleq \Delta^{-1}\mathcal{H}\mathcal{L}^*\mathcal{S}^{-*} = \mathcal{P}\mathcal{S}^{-*}$ has spectral radius strictly less than unity. Let us therefore identify the Hankel operator of $\mathcal{T}$. Thus,

$$\left[\begin{array}{c|c} \mathcal{T}_- & \mathcal{T}_A \\ \hline \mathcal{T}_H & \mathcal{T}_+ \end{array}\right] = \left[\begin{array}{c|c} \mathcal{P}_- & \mathcal{P}_A \\ \hline \mathcal{P}_H & \mathcal{P}_+ \end{array}\right] \left[\begin{array}{c|c} \mathcal{S}_-^{-*} & -\mathcal{S}_-^{-*}\mathcal{S}_H^*\mathcal{S}_+^{-*} \\ \hline 0 & \mathcal{S}_+^{-*} \end{array}\right],$$

which implies

$$\mathcal{T}_H = \mathcal{P}_H \mathcal{S}_-^{-*}.$$

Now $\gamma > \gamma_c$, if, and only if, $\mathcal{T}_H^* \mathcal{T}_H < I$, or equivalently, if, and only if, $\mathcal{S}_-^{-1} \mathcal{P}_H^* \mathcal{P}_H \mathcal{S}_-^{-*} < I$. But this implies

$$\mathcal{S}_- \mathcal{S}_-^* > \mathcal{P}_H^* \mathcal{P}_H.$$

Finally, since $\mathcal{S}\mathcal{S}^* = \gamma^2 I - \mathcal{E}$, we have $\mathcal{S}_- \mathcal{S}_-^* = \gamma^2 I - \mathcal{E}_-$, so that $\gamma > \gamma_c$, if, and only if,

$$\gamma^2 I > \mathcal{E}_- + \mathcal{P}_H^* \mathcal{P}_H.$$

But this readily implies $\gamma_c^2 = \|\mathcal{E}_- + \mathcal{P}_H^* \mathcal{P}_H\|_\infty$.

■

4 Some Special Cases

In this section we shall use the result of Theorem 2 to compute γ_c for the two important special cases of equalization (or its dual problem tracking) and filtering signals from additive noise.

4.1 The Equalization Problem

In the equalization problem we have $\mathcal{L} = I$. Referring back to Fig. 1, this means that we would like to estimate the unknown input signal $\{u_i\}$ from noisy measurements of the output of the linear time-invariant system $\mathcal{H}$. In this sense, the causal equalization problem can be regarded as the problem of causally inverting a linear system in the presence of additive noise.

In what follows, we shall assume that the linear system $\mathcal{H}$ is square, *i.e.*, that its impulse response $\{H_i\}$ are elements of $\mathcal{C}^{m\times m}$. For the more general case, see [6].

Theorem 4 (H^∞ Equalization) *Consider the causal Laurent operator $\mathcal{H}$, where the impulse response $\{H_i\}$ is a sequence of elements of $\mathcal{C}^{m\times m}$, and suppose we want to find γ_c, where*

$$\gamma_c = \inf_{\text{causal}\mathcal{K}} \left\| \begin{bmatrix} I - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty .$$

(i) If $\mathcal{H}$ is minimum phase, i.e., *if $\mathcal{H}^{-1}$ is causal, then*

$$\gamma_c^2 = \gamma_s^2 = \sup_{\omega\in[0,2\pi]} \bar{\sigma}\left[\left(I + H^*(e^{j\omega})H(e^{j\omega})\right)^{-1}\right]. \tag{27}$$

(ii) If $\mathcal{H}$ is non-minimum phase, i.e., *if $\mathcal{H}^{-1}$ is non-causal, then*

$$\gamma_c^2 = 1. \tag{28}$$

In other words, if the system $\mathcal{H}$ has a causal inverse then causal equalizers offer the same performance as non-causal equalizers that have access to future observations. On the other hand, if $\mathcal{H}$ does not have a causal inverse then causal equalization is not possible from the H^∞ point of view, since $\gamma = 1$ is the disturbance attenuation obtained by performing no equalization at all! Indeed $\mathcal{K} = 0$ yields

$$\mathcal{T}_{\mathcal{K}} = \begin{bmatrix} I & 0 \end{bmatrix}, \tag{29}$$

so that $\|\mathcal{T}_{\mathcal{K}}\|_\infty = 1$.

Proof of Theorem 4: To prove part (i) assume that $\mathcal{H}$ is minimum phase. We will show that

$$\gamma > \gamma_s \;\Rightarrow\; \gamma > \gamma_c$$

which establishes $\gamma_s \geq \gamma_c$. But since we already know that $\gamma_s \leq \gamma_c$ this readily implies that $\gamma_s = \gamma_c$.

Now since $\mathcal{H}$ is minimum phase, $H(z)$ has no unit circle zeros and hence

$$\gamma_s^2 = \sup_{\omega\in[0,2\pi]} \bar{\sigma}\left[\left(I + H^*(e^{j\omega})H(e^{j\omega})\right)^{-1}\right] < 1.$$

Choose now a γ such that $1 > \gamma > \gamma_s$. We thus have the following series of arguments:

$$\begin{aligned} 1 > \gamma > \gamma_s \quad &\Rightarrow \quad \gamma^2 I > (I + \mathcal{H}^*\mathcal{H})^{-1} \\ &\Rightarrow \quad (\gamma^{-2} - 1)I < \mathcal{H}^*\mathcal{H} \\ &\Rightarrow \quad \frac{1}{\gamma^{-2} - 1} I > \mathcal{H}^{-1}\mathcal{H}^{-*} \end{aligned}$$

$$\Rightarrow \frac{1}{\gamma^{-2}-1} I > \left[\begin{array}{c|c} \mathcal{H}_-^{-1} & 0 \\ \hline -\mathcal{H}_2^{-1}\mathcal{H}_H\mathcal{H}_-^{-1} & \mathcal{H}_2^{-1} \end{array}\right] \left[\begin{array}{c|c} \mathcal{H}_-^{-*} & \times \\ \hline 0 & \times \end{array}\right]$$

$$\Rightarrow \frac{1}{\gamma^{-2}-1} \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}\right] > \left[\begin{array}{c|c} \mathcal{H}_-^{-1}\mathcal{H}_-^{-*} & \times \\ \hline \times & \times \end{array}\right]$$

$$\Rightarrow \frac{1}{\gamma^{-2}-1} I > \mathcal{H}_-^{-1}\mathcal{H}_-^{-*},$$

where $\times$ denotes irrelevant entries. We can now unwind this last expression to get $\gamma^2 I > (I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}$. But this implies that $\gamma > \gamma_c$, and we are done. [Note that the key step in the above sequence of arguments was the fourth step where we used the fact that $\mathcal{H}^{-1}$ is causal, and hence lower triangular. Our argument would not follow through if $\mathcal{H}^{-1}$ were otherwise.]

To prove part (ii) assume now that $\mathcal{H}^{-1}$ is non-causal. This implies that $H(z)$ has a zero outside the unit circle. Let p ($|p| > 1$) be such a non-minimum phase zero of $H(z)$. Then

$$\underbrace{\begin{bmatrix} \ddots & & & \\ \ddots & H_0 & & \\ \dots & H_1 & H_0 & \\ \dots & H_2 & H_1 & H_0 \end{bmatrix}}_{\mathcal{H}_-} \underbrace{\begin{bmatrix} \vdots \\ p^{-2} \\ p^{-1} \\ 1 \end{bmatrix}}_{X} = H(p) \begin{bmatrix} \vdots \\ p^{-2} \\ p^{-1} \\ 1 \end{bmatrix} = 0. \tag{30}$$

Note that $X \in l_-^2$ since $|p| > 1$. The above equation shows that X is an eigenvector of $\mathcal{H}_-$ with eigenvalue $H(p)$. Now

$$(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1} X = \left(I - \mathcal{H}_-^*(I + \mathcal{H}_-\mathcal{H}_-^*)^{-1}\mathcal{H}_-\right) X = X,$$

so that $(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}$ has an eigenvalue of unity. Since $(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1} \leq I$ this implies that

$$\left\| (I + \mathcal{H}_-^*\mathcal{H}_-)^{-1} \right\|_\infty = 1,$$

which is our desired result. [Note that the key to this derivation was the fact that $H(z)$ had a non-minimum phase zero, *i.e.*, $|p| > 1$. The argument is no longer valid if $H(z)$ has only minimum phase zeros, since the vectors X that correspond to such zeros with $|p| < 1$ do not belong to l_-^2.]

$\blacksquare$

We finally should remark that the result of Theorem 4 follows with relative ease from Eq. (25). It would have been much more difficult to conceive and prove using Eq. (26).

4.2 Filtering Signals from Additive Noise

In this problem we have $\mathcal{L} = \mathcal{H}$. Referring back to Fig. 1, this means that we would like to estimate the unknown signal $\{s_i\}$ from the "signal-plus-noise" observations $\{y_i = s_i + v_i\}$. The result is given below. For further details see [7].

Theorem 5 (H^∞ Filtering of Signals in Additive Noise) *Consider the causal Laurent operator $\mathcal{H}$, and suppose we want to find γ_c, where*

$$\gamma_c = \inf_{\text{causal}\mathcal{K}} \left\| \begin{bmatrix} \mathcal{H} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty .$$

Then we have

$$\gamma_c^2 = \gamma_s^2 = \sup_{\omega \in [0,2\pi]} \bar{\sigma} \left[H(e^{j\omega}) \left(I + H^*(e^{j\omega}) H(e^{j\omega}) \right)^{-1} H^*(e^{j\omega}) \right]. \tag{31}$$

Note that the above result shows that, from an H^∞ point of view, causal estimators have the same performance as non-causal estimators when filtering signals from additive noise.

Proof of Theorem 5: We will show that

$$\gamma > \gamma_s \;\Rightarrow\; \gamma > \gamma_c$$

which establishes $\gamma_s \geq \gamma_c$. But since we already know that $\gamma_s \leq \gamma_c$ this readily implies that $\gamma_s = \gamma_c$.

First note that since $H(z)$ is analytic on $|z| = 1$, $H(e^{j\omega})$ is bounded for all $\omega \in [0, 2\pi]$, and hence

$$\gamma_s^2 = \sup_{\omega \in [0,2\pi]} \bar{\sigma} \left[H(e^{j\omega}) \left(I + H^*(e^{j\omega}) H(e^{j\omega}) \right)^{-1} H^*(e^{j\omega}) \right] < 1.$$

Choose now a γ such that $1 > \gamma > \gamma_s$. We thus have the following series of arguments:

$$\begin{aligned}
1 > \gamma > \gamma_s \quad & \Rightarrow \quad \gamma^2 I > \mathcal{H}(I + \mathcal{H}^*\mathcal{H})^{-1}\mathcal{H}^* = I - (I + \mathcal{H}\mathcal{H}^*)^{-1} \\
& \Rightarrow \quad (I + \mathcal{H}\mathcal{H}^*)^{-1} > (1 - \gamma^2) I \\
& \Rightarrow \quad I + \mathcal{H}\mathcal{H}^* < \frac{1}{1-\gamma^2} I \\
& \Rightarrow \quad \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}\right] + \left[\begin{array}{c|c} \mathcal{H}_- & 0 \\ \hline \mathcal{H}_H & \mathcal{H}_+ \end{array}\right] \left[\begin{array}{c|c} \mathcal{H}_-^* & \mathcal{H}_H^* \\ \hline 0 & \mathcal{H}_+^* \end{array}\right] < \frac{1}{1-\gamma^2} \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}\right] \\
& \Rightarrow \quad \left[\begin{array}{c|c} I + \mathcal{H}_-\mathcal{H}_-^* & \times \\ \hline \times & \times \end{array}\right] < \frac{1}{1-\gamma^2} \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}\right] \\
& \Rightarrow \quad I + \mathcal{H}_-\mathcal{H}_-^* < \frac{1}{1-\gamma^2} I,
\end{aligned}$$

where $\times$ denotes irrelevant entries. We can now unwind this last expression to get $\gamma^2 I > \mathcal{H}_-(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{H}_-^*$. But this implies that $\gamma > \gamma_c$, and we are done.

■

Note, once more, that the result of Theorem 5 would have been considerably more to difficult to conceive and prove using Eq. (25).

5 Worst-Case Non-Estimability and Worst-Case Complete Estimability

As mentioned earlier, explicit frequency-domain expressions for γ_c are currently not available. However, in general one can always claim that

$$\gamma_s \leq \gamma_c \leq \gamma_n \triangleq \|\mathcal{L}\|_\infty , \tag{32}$$

where the upper bound follows from the fact that γ_n corresponds to performing no estimation, *i.e.*, $\mathcal{K} = 0$, since for this choice of $\mathcal{K}$ we have $\mathcal{T}_{\mathcal{K}} = \begin{bmatrix} \mathcal{L} & 0 \end{bmatrix}$.

Therefore a natural question to ask is what are the conditions on $\mathcal{L}$ and $\mathcal{H}$ for γ_c to achieve either of the above upper and lower bounds? This is also an important question since:

- $\gamma_c = \gamma_s$ corresponds to an *easy* estimation problem, since here causal estimators have the same performance as noncausal ones.

- $\gamma_c = \gamma_n$ corresponds to a *difficult* estimation problem, since here causal estimators cannot offer any improvement over not estimating at all ($\mathcal{K} = 0$).

Due to the importance of the above concepts, we shall call the pair $\{\mathcal{L}, \mathcal{H}\}$ *worst-case complete estimable* if $\gamma_c = \gamma_s$ and *worst-case non-estimable* if $\gamma_c = \gamma_n$.

From our analysis of the previous section we already know the answer to whether $\{\mathcal{L}, \mathcal{H}\}$ is worst-case complete estimable or worst-case non-estimable for the special cases of equalization and filtering signal from additive noise. For the equalization problem the answer depends on whether $\mathcal{H}$ is minimum phase or not: if $\mathcal{H}$ is minimum phase than $\{I, \mathcal{H}\}$ is worst-case complete estimable, and if $\mathcal{H}$ is non-minimum phase $\{I, \mathcal{H}\}$ is worst-case non-estimable. For the problem of filtering signal from additive noise, the pair $\{\mathcal{H}, \mathcal{H}\}$ is always worst-case complete estimable.

Unfortunately, we do not currently have a characterization of when the arbitrary pair $\{\mathcal{L}, \mathcal{H}\}$ is worst-case complete estimable. However, for worst-case non-estimability we do have such a characterization, as described next.

5.1 Worst-Case Non-Estimability

Theorem 6 (Worst-Case Non-Estimability) *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$ and define*

$$\gamma_s = \inf_{\mathcal{K}} \left\| \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty ,$$

and

$$\gamma_c = \inf_{\text{causal } K} \left\| \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty .$$

Suppose, moreover, that $\gamma_s < \gamma_n = \|\mathcal{L}\|_\infty$. Then we have the following:

(i) $\gamma_c = \gamma_n$, if, and only if, there exists an $x \in l^2_-$ such that

(a) $\|\mathcal{L}_- x\|_2 = \|\mathcal{L}\|_\infty \|x\|_2$.

(b) $\mathcal{H}_- x = 0$.

(ii) If the corresponding z-transforms $L(z)$ and $H(z)$ are scalar and rational, then $\gamma_c = \gamma_n$, if, and only if,

(a) $L(z)$ is all-pass, i.e., *$|L(e^{j\omega})| = constant$, for all $\omega \in [0, 2\pi]$.*

(b) The number of non-minimum phase zeros of $H(z)$ (counting multiplicities) is greater than the McMillan degree of $L(z)$.

The above theorem gives a very simple characterization of worst-case non-estimability. Conditions (i)-(a) and (i)-(b) are valid for general (nonrational and matrix-valued) $L(z)$ and $H(z)$. When $L(z)$ and $H(z)$ are scalar and rational these conditions simplify to (ii)-(a) and (ii)-(b). As seen, here non-estimability depends on three simple properties: the all-passness of $L(z)$, the number of non-minimum phase (outside the unit circle) zeros of $H(z)$, and the McMillan degree of $L(z)$. We should also mention that there is a slightly more involved characterization of non-estimability for the case of rational, but matrix-valued, $L(z)$ and $H(z)$ (which involves the condition $\bar{\sigma}\left[L(e^{j\omega})\right] = \text{constant}$) but for brevity we shall not give it here.

Proof of Theorem 6: We shall first prove that

$$\gamma_c^2 = \sup_{x \neq 0 \in l^2_-} \frac{\|\mathcal{L}_- x\|_2^2}{\|x\|_2^2 + \|\mathcal{H}_- x\|_2^2}. \tag{33}$$

To this end, note that $\gamma > \gamma_c$ if, and only if,

$$\begin{aligned}
\gamma^2 I - \mathcal{L}_-(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^* \quad &\Leftrightarrow \quad \begin{bmatrix} I + \mathcal{H}_-^* \mathcal{H}_- & \mathcal{L}_-^* \\ \mathcal{L}_- & \gamma^2 I \end{bmatrix} > 0 \\
&\Leftrightarrow \quad I + \mathcal{H}_-^* \mathcal{H}_- - \gamma^{-2} \mathcal{L}_-^* \mathcal{L}_- > 0 \\
&\Leftrightarrow \quad x^*(I + \mathcal{H}_-^* \mathcal{H}_- - \gamma^{-2} \mathcal{L}_-^* \mathcal{L}_-) x > 0, \quad \forall x \neq 0 \in l^2_- \\
&\Leftrightarrow \quad \|x\|_2^2 + \|\mathcal{H}_- x\|_2^2 - \gamma^{-2} \|\mathcal{L}_- x\|_2^2, \quad \forall x \neq 0 \in l^2_- \\
&\Leftrightarrow \quad \gamma^2 > \frac{\|\mathcal{L}_- x\|_2^2}{\|x\|_2^2 + \|\mathcal{H}_- x\|_2^2}, \quad \forall x \neq 0 \in l^2_-,
\end{aligned}$$

which yields the desired result (33). We should also mention that a similar argument shows that

$$\gamma_s^2 = \sup_{x \neq 0 \in l^2} \frac{\|\mathcal{L} x\|_2^2}{\|x\|_2^2 + \|\mathcal{H} x\|_2^2}. \tag{34}$$

Now since we have assumed $\gamma_s < \gamma_n$, we can only have $\gamma_c = \gamma_n$ if $\gamma_s < \gamma_c$. This therefore implies that the supremum in (33) is achievable by an $x \in l_-^2$. [If the supremum were not achievable in l_-^2, then γ_c would coincide with the γ_s given in (34).] Thus, $\gamma_c = \gamma_n$ if, and only if, there exists some $x \in l_-^2$ such that

$$\frac{\|\mathcal{L}_- x\|_2^2}{\|x\|_2^2 + \|\mathcal{H}_- x\|_2^2} = \gamma_n^2 = \|\mathcal{L}\|_\infty^2 \, ,$$

or, in other words, if, and only if,

$$\left(\|\mathcal{L}_- x\|_2^2 - \|\mathcal{L}\|_\infty^2 \, \|x\|_2^2 \right) - \|\mathcal{L}\|_\infty^2 \, \|\mathcal{H}_- x\|_2^2 = 0.$$

But since $\|\mathcal{L}_- x\|_2^2 - \|\mathcal{L}\|_\infty^2 \, \|x\|_2^2 \leq 0$, we conclude that the above equality can hold if, and only if,

$$\|\mathcal{L}_- x\|_2^2 = \|\mathcal{L}\|_\infty^2 \, \|x\|_2^2 \quad \text{and} \quad \|\mathcal{H}_- x\|_2^2 = 0$$

which yields precisely Conditions (i)-(a) and (i)-(b).

To prove the second claim assume that $L(z)$ and $H(z)$ are scalar rational transfer functions. We will first show that an $x \in l_-^2$ that achieves $\|\mathcal{L}\|_\infty$ exists if, and only if, $L(z)$ is all-pass. To this end, suppose that $L(z)$ is not all-pass so that $\|\mathcal{L}\|_\infty$ is achieved at only a finite number of frequencies (which correspond to the peak frequencies of $|L(e^{j\omega})|$). This means that $\|\mathcal{L}\|_\infty$ is achieved by signals that are the sum of a finite number of sinusoids (corresponding to these peaks). But since the sum of a finite number of non-zero sinusoidal signals can never belong to l_-^2 (or l^2, for that matter), $\|\mathcal{L}\|_\infty$ cannot be achieved by any $x \in l_-^2$. To prove the other direction, suppose that $L(z)$ is all-pass. Then if we consider the vector $\begin{bmatrix} x \\ 0 \end{bmatrix} \in l^2$, we have

$$\left\| \mathcal{L} \begin{bmatrix} x \\ 0 \end{bmatrix} \right\|_2^2 = \|\mathcal{L}\|_\infty^2 \left\| \begin{bmatrix} x \\ 0 \end{bmatrix} \right\|_2^2 ,$$

since $\|\mathcal{L}y\|_2^2 = \|\mathcal{L}\|_\infty^2 \|y\|_2^2$, for all $y \in l^2$. But the LHS is equal to $\left\| \begin{bmatrix} \mathcal{L}_- x \\ \mathcal{L}_H x \end{bmatrix} \right\|_2^2$ and the RHS equal to $\|\mathcal{L}\|_\infty^2 \|x\|_2^2$. We thus have

$$\|\mathcal{L}_- x\|_2^2 + \|\mathcal{L}_H x\|_2^2 = \|\mathcal{L}\|_\infty^2 \|x\|_2^2.$$

We can now conclude that $\|\mathcal{L}_- x\|_2^2 = \|\mathcal{L}\|_\infty^2 \|x\|_2^2$ if, and only if, $\mathcal{L}_H x = 0$. But since, in the rational case, $\mathcal{L}_H$ has finite rank (equal to the McMillan degree of $L(z)$) such an $x \in l_-^2$ can always be found. This establishes Condition (ii)-(a).

We thus conclude that $\gamma_c = \gamma_n$ if, and only if, there exists an $x \in l_-^2$ such that

$$\mathcal{L}_H x = 0 \quad \text{and} \quad \mathcal{H}_- x = 0.$$

We shall presently show that this is possible if, and only if, the number of non-minimum phase zeros of $H(z)$ is greater than the McMillan degree of $L(z)$. To do so, let us begin by identifying the null-space of

$$\mathcal{H}_- = \begin{bmatrix} \ddots & & & \\ \ddots & H_0 & & \\ \dots & H_1 & H_0 & \\ \dots & H_2 & H_1 & H_0 \end{bmatrix}.$$

It is straightforward to see that if $H(z)$ has non-minimum phase zeros $\{p_i, |p_i| > 1\}_{i=1}^n$, each with multiciplity $\{\lambda_i\}_{i=1}^n$, then the null-space of $\mathcal{H}_-$ is given by

$$\text{Span}\left\{ \begin{bmatrix} \vdots \\ \vdots \\ p_i^{-3} \\ p_i^{-2} \\ p_i^{-1} \\ 1 \end{bmatrix}, \begin{bmatrix} \vdots \\ \vdots \\ 3p_i^{-2} \\ 2p_i^{-1} \\ 1 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} \vdots \\ \lambda_i p_i^{-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\}_{i=1}^n.$$

For simplicity, we shall henceforth assume that $H(z)$ has no repeated roots. [Our arguments can be carried over to the general case with some notational care.] Thus the null-space of $\mathcal{H}_-$ is given by

$$\text{Span}\left\{ \begin{bmatrix} \vdots \\ p_1^{-2} \\ p_1^{-1} \\ 1 \end{bmatrix}, \begin{bmatrix} \vdots \\ p_2^{-2} \\ p_2^{-1} \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} \vdots \\ p_n^{-2} \\ p_n^{-1} \\ 1 \end{bmatrix} \right\}.$$

Let us now compute $\mathcal{L}_H x$ for every basis vector in the above null-space of $\mathcal{H}_-$. To this end, note that $\mathcal{L}_H$ can be written as

$$\mathcal{L}_H = \mathcal{O}_L \mathcal{C}_L,$$

where $\mathcal{O}_L$ and $\mathcal{C}_L$ are the full-rank observability and controllability matrices corresponding to $L(z)$:

$$\mathcal{O}_L^* = \begin{bmatrix} \dots & F^{*2}H^* & F^*H^* & H^* \end{bmatrix} \quad \text{and} \quad \mathcal{C}_L = \begin{bmatrix} \dots & F^2G & FG & G \end{bmatrix},$$

that have rank d, equal to the McMillan degree of $L(z)$. [Note here that F is a stable matrix.] Thus,

$$\mathcal{L}_H \begin{bmatrix} \vdots \\ p_i^{-2} \\ p_i^{-1} \\ 1 \end{bmatrix} = \mathcal{O}_L \begin{bmatrix} \dots & F^2G & FG & G \end{bmatrix} \begin{bmatrix} \vdots \\ p_i^{-2} \\ p_i^{-1} \\ 1 \end{bmatrix} = \mathcal{O}_L (I - p_i^{-1}F)^{-1}G,$$

where the inverse exists since p_i cannot be an eigenvalue of F, as F is stable. Since $\mathcal{O}_L$ is full-rank, there will exist an x in the null-space of $\mathcal{H}_-$, such that $\mathcal{L}_H x = 0$ if, and only if, the vectors

$$\left\{ (I - p_1^{-1}F)^{-1}G, (I - p_2^{-1}F)^{-1}G, \dots, (I - p_n^{-1}F)^{-1}G \right\},$$

are linearly dependent. Multiplying throughout by the nonsingular matrix $\prod_{i=1}^n (I - p_1^{-1}F)$ this is equivalent to the condition that the vectors

$$\left\{ \prod_{i \neq 1} (I - p_i^{-1}F)G, \prod_{i \neq 2} (I - p_i^{-1}F)G, \dots, \prod_{i \neq n} (I - p_i^{-1}F)G \right\}, \tag{35}$$

be linearly independent. But note that

$$\prod_{i \neq 1} (I - p_i^{-1}F)G = G - \left(\sum_{i \neq 1} p_i^{-1} \right) FG + \frac{1}{2} \left(\sum_{i \neq 1, i \neq j} p_i^{-1} p_j^{-1} \right) F^2G + \dots + (-1)^{n-1} \left(\prod_{i \neq 1} p_i^{-1} \right) F^{n-1}G,$$

so that we may write

$$\begin{bmatrix} \prod_{i \neq 1} (I - p_i^{-1}F)G & \prod_{i \neq 2} (I - p_i^{-1}F)G & \dots & \prod_{i \neq n} (I - p_i^{-1}F)G \end{bmatrix} =$$

$$\left[\begin{array}{ccccc} G & FG & F^2G & \ldots & F^{n-1}G \end{array}\right] \underbrace{\left[\begin{array}{cccc} 1 & 1 & \ldots & 1 \\ -\sum_{i\neq 1} p_i^{-1} & -\sum_{i\neq 2} p_i^{-1} & \ldots & -\sum_{i\neq n} p_i^{-1} \\ \frac{1}{2}\sum_{i\neq 1, i\neq j} p_i^{-1}p_j^{-1} & \frac{1}{2}\sum_{i\neq 2, i\neq j} p_i^{-1}p_j^{-1} & \ldots & \frac{1}{2}\sum_{i\neq n, i\neq j} p_i^{-1}p_j^{-1} \\ \vdots & \vdots & \ldots & \vdots \\ (-1)^{n-1}\prod_{i\neq 1} p_i^{-1} & (-1)^{n-1}\prod_{i\neq 2} p_i^{-1} & \ldots & (-1)^{n-1}\prod_{i\neq n} p_i^{-1} \end{array}\right]}_{\triangleq M}.$$

Using the symmetries of M with respect to the p_i, it can be shown that

$$\det\ M = \sqrt{\left|\prod_{i\neq j}(p_i^{-1} - p_j^{-1})\right|}.$$

Thus, since the $\{p_i\}$ are distinct, the collection of vectors in (35) will be linearly dependent if, and only if, the matrix

$$\left[\begin{array}{ccccc} G & FG & F^2G & \ldots & F^{n-1}G \end{array}\right],$$

has linearly dependent rows. But since $\{F, G\}$ is controllable (we have taken $\mathcal{C}_L$ to be full rank), this is true if, and only if, $n > d$. Since n represents the number of non-minimum phase zeros of $H(z)$ and d represents the McMillan degree of $L(z)$, we have our desired result.

■

Notet that the above result confirms the result of Theorem 4, part (ii), on equalization. Indeed in the equalization problem $L(z) = 1$ is clearly all-pass and has McMillan degree zero. Thus if $H(z)$ has any non-minimum phase zero, the equalization problem is worst-case non-estimable and $\gamma_c = \gamma_n = 1$.

5.2 Application to Equalization with Delay

Theorem 6 gives a very simple characterization of worst-case non-estimability. This is quite useful since it will allow us to recognize the occurrence of worst-case non-estimability in various applications, and to design estimation (and control) scenarios to avoid it. We shall presently demonstrate this by considering the problem of equalizing a scalar rational LTI system, $H(z)$.

Recall from Theorem 4 that if $H(z)$ were non-minimum phase then causal equalization is not possible since $\gamma_c = \gamma_n = 1$. One may then speculate whether it is possible to causally equalize $H(z)$ by allowing a finite amount of delay. In this case, one would like to estimate u_{i-d}, for some $d > 0$, using the observations $\{y_j, j \leq i\}$. Mathematically, this corresponds to choosing $L(z) = z^{-d}$ in the two-block problem of Sec. 2. [Clearly, $d = \infty$ corresponds to non-causal equalization.] The natural question to ask is what is the delay necessary to guarantee $\gamma_c < \gamma_n = 1$? The answer is given by the following lemma.

Lemma 1 (Equalization with Delay) *Consider the Laurent operator with scalar rational transfer function $H(z)$, and define*

$$\gamma_c = \inf_{\text{causal } K(\cdot)} \left\| \begin{bmatrix} z^{-d} - K(z)H(z) & -K(z) \end{bmatrix} \right\|_\infty .$$

Then if the number of non-minimum phase zeros of $H(z)$ is given by n, we have:

(i) $\gamma_c < 1$ *if* $d \geq n$.

(ii) $\gamma_c = 1$ *if* $d < n$.

Proof: The lemma clearly follows from Theorem 6, part (ii), since $L(z) = z^{-d}$ is all-pass and has McMillan degree d.

■

Therefore the minimum amount of delay is given by the number of non-minimum phase zeros of $H(z)$. Another interesting question would be determining the amount of delay required to guarantee $\gamma_c = \gamma_s$. But this is essentially the question of worst-case complete estimability for which we currently have no answer.

6 Conclusion

In this paper we obtained a new formula for the minimum disturbance attenuation in two block H^∞ problems. This new formula has the benefits of being much simpler than the "mixed Toeplitz-plus-Hankel spectral radius" formula currently available, and of being similar in structure to the formula for the the minimum disturbance attenuation in non-causal two block problems. [Essentially, the only difference is the replacement of Laurent operators by

Toeplitz ones.] Due to these benefits we were able to analyze in detail the behaviour of the minimum disturbance attenuation for the two important problems of equalization (or tracking, by duality) and filtering signals in additive noise. While for general estimation problems this is not yet possible, we did introduce the concepts of worst-case complete estimability, essentially when causal estimators have the same H^∞ performance as noncausal ones, and worst-case non-estimability, essentially when causal estimators cannot reduce the disturbance attenutation from their a priori values. We were also able to give a complete characterization of worst-case non-estimability and showed the value of this concept and characterization by studying the problem of equalization with a finite amount of delay. Finally, we should mention that open problems suggested by this paper include finding a characterization of worst-case complete estimability, studying the consequences of the results to four-block H^∞ problems, and, perhaps, finding more explicit frequency-domain characterizations of the minimum achievable disturbance attenuation.

References

[1] M. Verma and E.A. Jonckheere. l^∞ compensation with mixed sensitivity as a broadband matching problem. *Systems and Control Letters*, 4:125–129, 1984.

[2] A. Feintuch and B.A. Francis. Uniformly optimal control of linear systems. *Automatica*, 21:563–574, 1985.

[3] E.A. Jonckheere and M. Verma. A spectral characterization of H^∞ optimal feedback performance and its efficient computation. *Systems and Control Letters*, pages 124–131, 1986.

[4] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis. State-space solutions to standard H^2 and H^∞ control problems. *IEEE Trans. on Automatic Control*, 34:831–847, 1989.

[5] Z. Nehari. On bounded bilinear forms. *Annals of Mathematics*, 15(1):153–162, 1957.

[6] B. Hassibi, A.E. Erdogan, and T. Kailath. Equalization with an H^∞ criterion. *Submitted to IEEE Transactions on Information Theory*, 1997.

[7] B. Hassibi and T. Kailath. H^∞ filtering of signals in additive noise. *Submitted to IEEE Transactions on Signal Processing*, 1997.